Fast numerical design of spatial-selective rf pulses in MRI using Krotov and quasi-Newton based optimal control methods

Mads S. Vinding, Ivan I. Maximov, Zdeněk Tošner, and Niels Chr. Nielsen

---

Articles you may be interested in

Stabilized quasi-Newton optimization of noisy potential energy surfaces
J. Chem. Phys. **142**, 034112 (2015); 10.1063/1.4905665

A variable echo-number method for estimating R 2 in MRI-based polymer gel dosimetry
Med. Phys. **38**, 975 (2011); 10.1118/1.3544659

Quasi-Newton parallel geometry optimization methods
J. Chem. Phys. **133**, 034116 (2010); 10.1063/1.3455719

MRI-compatible transurethral ultrasound system for the treatment of localized prostate cancer using rotational control
Med. Phys. **35**, 1346 (2008); 10.1118/1.2841937

Magnetically-assisted remote control (MARC) steering of endovascular catheters for interventional MRI: A model for deflection and design implications
Med. Phys. **34**, 3135 (2007); 10.1118/1.2750963

Fast numerical design of spatial-selective rf pulses in MRI using Krotov and quasi-Newton based optimal control methods

Mads S. Vinding,[1] Ivan I. Maximov,[2] Zdeněk Tošner,[3] and Niels Chr. Nielsen[1,a)]

[1]*Center for Insoluble Protein Structures (inSPIN), Interdisciplinary Nanoscience Center (iNANO) and Department of Chemistry, Aarhus University, Langelandsgade 140, DK-8000 Aarhus C, Denmark*
[2]*Institute of Neuroscience and Medicine 4, Forschungszentrum Jülich GmbH, Jülich, Germany*
[3]*Department of Chemistry, Faculty of Science, Charles University in Prague, Hlavova 8, CZ-128 43, Czech Republic*

The use of increasingly strong magnetic fields in magnetic resonance imaging (MRI) improves sensitivity, susceptibility contrast, and spatial or spectral resolution for functional and localized spectroscopic imaging applications. However, along with these benefits come the challenges of increasing static field (B_0) and rf field (B_1) inhomogeneities induced by radial field susceptibility differences and poorer dielectric properties of objects in the scanner. Increasing fields also impose the need for rf irradiation at higher frequencies which may lead to elevated patient energy absorption, eventually posing a safety risk. These reasons have motivated the use of multidimensional rf pulses and parallel rf transmission, and their combination with tailoring of rf pulses for fast and low-power rf performance. For the latter application, analytical and approximate solutions are well-established in linear regimes, however, with increasing nonlinearities and constraints on the rf pulses, numerical iterative methods become attractive. Among such procedures, optimal control methods have recently demonstrated great potential. Here, we present a Krotov-based optimal control approach which as compared to earlier approaches provides very fast, monotonic convergence even without educated initial guesses. This is essential for *in vivo* MRI applications. The method is compared to a second-order gradient ascent method relying on the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton method, and a hybrid scheme Krotov-BFGS is also introduced in this study. These optimal control approaches are demonstrated by the design of a 2D spatial selective rf pulse exciting the letters "JCP" in a water phantom. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4739755]

I. INTRODUCTION

Beyond discussion, magnetic resonance imaging (MRI) is one of the strongest medical imaging modalities ever invented. Not only the anatomy and pathological tissue are visualizable, but also the function,[1] metabolism,[2] and microscopic structure[3–7] can be studied with MRI. This versatile modality with numerous medical and basic science applications is, however, suffering from low signal-to-noise ratio (SNR), and manufacturers of MRI systems aim at increasingly high field strengths to increase SNR, functional sensitivity, as well as spatial and spectral resolution. Indeed, a number of studies have shown that higher fields improve image quality,[8,9] localized spectroscopy (MRS),[10] and functional imaging (fMRI).[11]

Along with the attractive features of high fields, however, comes a number of technical challenges. The dielectric properties of samples at high field may interfere with radio frequency (rf) irradiation leading to spatially varying excitation patterns and thereby non-uniform image contrast. Such distortions are sample/patient specific and advanced coil setups can only mitigate the inhomogeneous B_1-field to a certain degree. Likewise, despite blood-oxygenation-level-dependent (BOLD) fMRI rely on T_2^*-contrast that is enhanced by increased field strength, B_0-inhomogeneity can become too severe near air cavities and bone structures preventing BOLD fMRI studies in such regions. Furthermore, with increasing rf frequencies, the energy deposition in the sample/patient also grows. Specific absorption rate (SAR) limitations placed by governmental administrations actually limit ultra-high fields from clinical use.

Multidimensional rf (MDRF) pulses have found renewed interest after the invention of parallel transmit (pTx) systems.[12,13] While MDRF pulses are capable of alleviating B_0- and B_1-inhomogeneities, they are quite long on single-transmit (sTx) systems relative to ordinary pulses. With pTx systems, the pulses can be accelerated as many-fold as there are separate coils in the transmit arrays. Such pulses tailored to the specific sample/patient may therefore have significant impact on the applicability of ultra-high-field MRI. Several well-established approaches exist for the design of MDRF pulses. The small-tip-angle (STA) approach introduced by Pauly *et al.*[14] provide an analytical foundation for such developments. Later, the same authors extended the formalism to include large tip-angles by restricting the underlying *k*-space trajectories to follow certain rules. This linear Fourier relationship between the excitation profile and the weighted *k*-space trajectory has an appreciated simplicity facilitating direct understanding of the function of the pulses. However, as experimental challenges are brought into play, the pulse

---

a)Author to whom correspondence should be addressed. Electronic mail: ncn@inano.au.dk.

design may more conveniently be handled by numerical, iterative optimization schemes.[15,16]

Optimal control (OC) theory[17–19] has lately found increasing use in magnetic resonance to handle dynamic systems with a large number of controls. Recent examples include pulse sequence design in liquid-[20,21] and solid-state[22,23] NMR, hyper-polarized NMR,[24,25] and MRI.[26–28] In these areas, OC has demonstrated a great potential to establish experiments adapting well to technical challenges and constraints included in the cost functional, as for example, rf power limitations,[27] rf inhomogeneity,[22,23] and rf envelope "jaggedness."[29,30] Addressing MRI, OC procedures have so far involved gradient-based methods[20,26,27,31] exploiting conjugated gradient optimization.[32] Such methods are known to be robust and convergent within a finite number of iterations. However, the gradient-based methods require one forward and one backward propagation in time for each iteration in addition to a number of forward propagations to evaluate the optimal direction and step-size. Optimization with gradients alone is also prone to convergence towards local optima and may need both plenty of iterations to reach the optimum and may need multiple repetitions to reach a satisfactory optimum without an educated initial starting point. In order to improve these cumbersome steps, Grissom *et al.*[31] have introduced fast optimal control procedures that exploit the nonuniform fast Fourier transform (NUFFT) approximation to the Bloch equation and they demonstrate significant reduction in computation time in the design of pTx MDRF pulses. However, the gradient-based algorithms are still complicated for practical realization and too time consuming for clinical applications.

In this work, we introduce a novel Krotov-based[19] monotonically convergent OC approach for MDRF pulse design. This method proves significantly faster than gradient-based methods. Without an educated initial guess, the Krotov approach convergences monotonically towards the global optimum. The Krotov approach has lately found applications in NMR spectroscopy,[24,29] quantum chemistry,[33,34] quantum control,[35] and dynamic nuclear polarization (DNP).[24] Furthermore, for comparison, we adopted a second-order gradient ascent quasi-Newton algorithm known as Broyden[36]-Fletcher[37]-Goldfarb[38]-Shanno.[39] Addressing NMR applications, Fouquieres *et al.*[40] recently demonstrated that BFGS has a robust and relatively steep convergence in the later iterations, which may complement the Krotov approach being significantly faster in the first iterations.[24,29] Adapting this strategy, Eitan *et al.*[41] found it worthwhile to combine Krotov with BFGS in relation to quantum control problems. In this work, we have adapted the Krotov-BFGS (multiple control variation of Eitan's approach) for the design of MDRF pulses. The overall aim is to present a comparative study of the three algorithms in relation to MRI applications.

The paper is organized as follows: In the theory section, we introduce key aspects of optimal control and MDRF pulse design. For the various optimal control approaches, we extend the formalism with step-wise procedures for numerical implementation. For more details, we refer the reader to a more extensive account in supplementary material.[42] In the results section, we compare the optimal methods with attention to the design of MDRF pulses within MRI applications. The power of optimal control design of MDRF pulses is demonstrated by excitation of the symbols "JCP" in a water phantom and excitation of the mouse brain with experimental verifications obtained on a 16.4 T vertical bore scanner.

II. THEORY

A. The optimal control formulation

The dynamics of the magnetization vector $\mathbf{M}(\mathbf{r}, t) = [M_x(\mathbf{r}, t), M_y(\mathbf{r}, t), M_z(\mathbf{r}, t)]^\top$ is governed by the Bloch equation, which when neglecting relaxation may be cast as

$$\dot{\mathbf{M}}(\mathbf{r}, t) = \mathbf{\Omega}(\mathbf{r}, t)\mathbf{M}(\mathbf{r}, t) \quad (1)$$

with

$$\mathbf{\Omega}(\mathbf{r}, t) = \begin{bmatrix} 0 & \gamma\mathbf{G}(t) \cdot \mathbf{r} + \Delta\omega_0(\mathbf{r}) & -\beta(\mathbf{r})\omega_y(t) \\ -\gamma\mathbf{G}(t) \cdot \mathbf{r} - \Delta\omega_0(\mathbf{r}) & 0 & \beta(\mathbf{r})\omega_x(t) \\ \beta(\mathbf{r})\omega_y(t) & -\beta(\mathbf{r})\omega_x(t) & 0 \end{bmatrix}. \quad (2)$$

Here the spins at position $\mathbf{r} = [x, y, z]^\top$ are assumed to be on resonance with respect to the applied rf pulse waveform $\gamma B_1(t) = \omega_x(t) + i\omega_y(t)$ and simultaneously subject to a time-varying field gradient $\mathbf{G}(t) = [G_x(t), G_y(t), G_z(t)]^\top$. The factor $\beta(\mathbf{r})$ corresponds to the spatial dependent B_1 inhomogeneity. This factor is measured as a B_1 map.[43] Off-resonance conditions, i.e., B_0 inhomogeneity, can be included by the $\Delta\omega_0(\mathbf{r})$ term; measured in a B_0 map.[44] Under ideal circumstances, $\beta(\mathbf{r})$ and $\Delta\omega_0(\mathbf{r})$ can be set equal to 1 and 0, respectively. The task is to invert the Bloch equation to derive the $B_1(t)$ waveform that for each position

$\mathbf{r}$ produces a desired magnetization $\mathbf{M}_{des}(\mathbf{r}) = [M_{x,des}(\mathbf{r}), M_{y,des}(\mathbf{r}), M_{z,des}(\mathbf{r})]^\top$ at the end of the period T.

In OC terms, this amounts to maximizing a cost functional given by

$$J(\mathbf{r}) = \Phi(\mathbf{M}(\mathbf{r}, T)) + \int_0^T \Xi(\mathbf{M}(\mathbf{r}, t), \omega_x(t), \omega_y(t), t)dt, \quad (3)$$

with $\Phi(\mathbf{M}(\mathbf{r}, T))$ representing the final cost (an efficiency functional dependent on the final magnetization, and possibly on the final time) and $\Xi(\mathbf{M}(\mathbf{r}, t), \omega_x(t), \omega_y(t), t)$ the running cost (a regularizable penalty functional). Both can be

accustomed to the specific problem. For example, to design an excitation pulse with high excitation fidelity for a fixed pulse duration, still, limited in rf power, the two functionals could be written as

$$\Phi(\mathbf{M}(\mathbf{r}, T)) = \frac{1}{M_0^2}\mathbf{M}(\mathbf{r}, T)^{\mathsf{T}} \cdot \mathbf{M}_{des}(\mathbf{r}), \qquad (4)$$

$$\Xi(\omega_x(t), \omega_y(t)) = -\lambda\left(\omega_x^2(t) + \omega_y^2(t)\right), \qquad (5)$$

where λ is a non-negative Tikhonov regularization parameter and M_0 is the magnitude of the equilibrium magnetization. The initial magnetization is in this case given by $\mathbf{M}(\mathbf{r}, t) = [0, 0, M_0]^{\mathsf{T}}$. The dependence of the spatial variable $\mathbf{r}$ in the cost functional $J(\mathbf{r})$ emphasizes the need for an integration over the entire space to yield a single measure of J. This representation of the functionals in Eqs. (4) and (5) will be used throughout this paper.

Pontryagin *et al.*[17] tackled the optimization problem by drawing on principles from calculus of variations and the Hamilton-Jacobi-Bellman equation. The cost functional in Eq. (3) is adjoined to Eq. (1) by introduction of Lagrange multipliers $L(\mathbf{r}, t)$ that yields an augmented cost functional

$$\bar{J}(\mathbf{r}) = \Phi(\mathbf{M}(\mathbf{r}, T)) + \int_0^T \Xi(\omega_x(t), \omega_y(t))dt$$

$$+ \int_0^T \mathbf{L}^{\mathsf{T}}(\mathbf{r}, t)\left(\mathbf{\Omega}(\mathbf{r}, t)\mathbf{M}(\mathbf{r}, t) - \dot{\mathbf{M}}(\mathbf{r}, t)\right)dt. \quad (6)$$

The problem may be formulated in terms of a Hamiltonian

$$\mathcal{H} = \Xi(\omega_x(t), \omega_y(t)) + \mathbf{L}^{\mathsf{T}}(\mathbf{r}, t)\mathbf{\Omega}(\mathbf{r}, t)\mathbf{M}(\mathbf{r}, t), \qquad (7)$$

which will become handy in the interpretation of the optimal control. Calculus of variations states that under optimal conditions, denoted with an *, four necessary conditions must be fulfilled

$$\frac{\delta\mathcal{H}^*}{\delta\mathbf{L}(\mathbf{r}, t)} = \dot{\mathbf{M}}^*(\mathbf{r}, t), \qquad (8)$$

$$\frac{\delta\mathcal{H}^*}{\delta\mathbf{M}(\mathbf{r}, t)} = -\dot{\mathbf{L}}^*(\mathbf{r}, t), \qquad (9)$$

$$\frac{\delta\Phi}{\delta\mathbf{M}(\mathbf{r}, T)} - \mathbf{L}^*(\mathbf{r}, T) = 0, \qquad (10)$$

$$\frac{\delta\mathcal{H}^*}{\delta\omega_\alpha} = 0, \ \alpha = x, y. \qquad (11)$$

The necessary conditions in Eqs. (8)–(11) provide tools for deriving manageable formulas that can be applied to our physical problem. For example, the condition in Eq. (9) leads to

$$\dot{\mathbf{L}}(\mathbf{r}, t) = \mathbf{\Omega}(\mathbf{r}, t)\mathbf{L}(\mathbf{r}, t), \qquad (12)$$

which is the equation of motion for the Lagrange multiplier. The magnetization, bounded by the initial state and the third condition in Eq. (10), provides us with a boundary for the Lagrange multiplier

$$\mathbf{L}(\mathbf{r}, T) = \mathbf{M}_{des}(\mathbf{r}), \qquad (13)$$

which happens to be the desired magnetization. Finally, Eq. (11) can be regarded as our gradient expression with respect to our controls

$$\frac{\delta\mathcal{H}^*}{\delta\omega_\alpha} \equiv \nabla J_\alpha$$

$$= -2\lambda\omega_\alpha(t) + \int_{\mathbf{r}} \beta(\mathbf{r})\mathbf{L}(\mathbf{r}, t) \times \mathbf{M}(\mathbf{r}, t)|_\alpha \, d\mathbf{r}, \quad (14)$$

explicitly from the cross-products we use the x and y components only, respectively, as signified by $\alpha = x, y$.

B. Discretization

Equation (1) can be solved as

$$\mathbf{M}(\mathbf{r}, t) = \hat{T}e^{\int_0^t \mathbf{\Omega}(\mathbf{r}, t')dt'}\mathbf{M}(\mathbf{r}, 0), \qquad (15)$$

where $\hat{T}$ is the Dyson time-ordering operator. For numerical optimizations, the formula and continuous variables described above need to be discretized and implemented into a framework for efficient calculations. In this work, we use MATLAB (Ref. 45) as a widely accessible platform for our optimizations.

The magnetization and Lagrange vectors will move along a trajectory during the discretized pulse. We define the optimal control pulse as an entity with N subpulses of duration $\Delta t = T/N$. The spatial geometry (space) in consideration for the design of rf pulses exciting a certain 2D area (throughout denoted as a MDRF pulse) is a plane consisting of P individual locations in which we demand a specific excitation profile. We denote certain positions in space and time using the variables $p \in \{1, \ldots, P\}$ and $n \in \{1, \ldots, N\}$, respectively. With N subpulses indexed as $\omega_{x, n}$ and $\omega_{y, n}$, there will be $N + 1$ magnetization states for a given position $\mathbf{r}_p$, as denoted $\mathbf{M}_{p,0}, \ldots, \mathbf{M}_{p,N}$. We assume the spins in one position are uncoupled from spins at any other position. Thus, the magnetization (and the Lagrange state) at a point p is propagated from one step to the next as

$$\mathbf{M}_{p,n} = \mathbf{U}_{p,n}\mathbf{M}_{p,n-1}, \qquad (16)$$

with $\mathbf{U}_{p,n}$ representing a standard rotation matrix, e.g., an axis-angle type or the matrix exponential of Eq. (2),

$$\mathbf{U}_{p,n} = e^{\mathbf{\Omega}_{p,n}\Delta t}, \qquad (17)$$

where $\mathbf{\Omega}_{p,n}$ is

$$\mathbf{\Omega}_{p,n} = \begin{bmatrix} 0 & \gamma\mathbf{G}_n \cdot \mathbf{r}_p + \Delta\omega_{0,p} & -\beta_p\omega_{y,n} \\ -\gamma\mathbf{G}_n \cdot \mathbf{r}_p - \Delta\omega_{0,p} & 0 & \beta_p\omega_{x,n} \\ \beta_p\omega_{y,n} & -\beta_p\omega_{x,n} & 0 \end{bmatrix}, \qquad (18)$$

with $\mathbf{G}_n$ the discretized form of $\mathbf{G}(t)$. Preferably, $\Delta t \ll T$ for accuracy implying a compromise between accuracy and a tolerable number of controls N.

For MDRF pulses, the desired magnetization may for example be encoded in a grid of, say, 64×64 points, i.e., $P = 4096$. As proposed by Conolly *et al.*,[26] it may be practically advantageous to arrange the individual magnetization and Lagrange vectors for all p in a stacked super-vector,

$$\mathbf{M}_n = \begin{bmatrix} \mathbf{M}_{1,n} \\ \mathbf{M}_{2,n} \\ \vdots \\ \mathbf{M}_{p,n} \\ \vdots \\ \mathbf{M}_{P,n} \end{bmatrix} \tag{19}$$

and the rotation matrices in a block-diagonal matrix,

$$\mathbf{U}_n = \begin{bmatrix} \ddots & & \\ & \mathbf{U}_{p,n} & \\ & & \ddots \end{bmatrix}. \tag{20}$$

In this manner, tedious loops over all spatial points can be avoided and the rotation of all spatial points may be represented by a single matrix vector product. However, we have found a computationally more efficient method to do the rotations and the cross-products. The details can be seen in the supplementary material.[42] For simplicity, we continue the theoretical details with the notation of stacked vectors and block-diagonal matrices.

C. Gradient ascent

The most widely used approach for optimal control in magnetic resonance has been the gradient-based method such as the gradient ascent pulse engineering (GRAPE) proposed by Khaneja and co-workers.[20] This method relies on the first-order gradient approximation derived in Sec. II A. For the kth iteration this method involves the following steps:

1. Propagate $\mathbf{M}_0^{(k)}$ forward to $\mathbf{M}_N^{(k)}$ using Eq. (1) and $\omega_{\alpha,n}^{(k)}$.
2. Propagate $\mathbf{L}_N^{(k)}$ backward to $\mathbf{L}_0^{(k)}$ using Eq. (12) and $\omega_{\alpha,n}^{(k)}$.
3. Update all controls with Eq. (14),

$$\omega_{\alpha,n}^{(k+1)}$$
$$= \omega_{\alpha,n}^{(k)} + \varepsilon \left(-2\lambda\omega_{\alpha,n}^{(k)} + \frac{1}{PM_0^2}\sum_{p=1}^{P}\beta_p\mathbf{L}_{p,n}^{(k)}\times\mathbf{M}_{p,n-1}^{(k)}\big|_\alpha \right). \tag{21}$$

4. Increment k and go to step 1, unless convergence criterion for cost functional is met.

The ε is a scalar representing the step-size. The simplicity of the GRAPE algorithm is worth noticing. One forward propagation of the magnetization from the initial to some final state and one backward propagation of the Lagrange multiplier from the desired state to some initial state yield two

paths. The difference in the paths is used as a gradient to update all controls at once being a hallmark of OC. The GRAPE approach has found widespread application for optimization of experiments within NMR (Refs. [20]–[23], [46], and [47]) and MRI (Refs. [26], [27], [31], and [48]) as mediated for example by implementation[23] into software packages such as SIMPSON.[49] We will in Sec. IV provide a practical example where SIMPSON in this setup has been used to design a MDRF pulse that excites the brain of a mouse.

D. The Krotov-based approach

While the GRAPE algorithm suggested a control update of the kind in Eq. (21), the Krotov algorithm uses all available information in the iteration cycle to update the controls using a cross-product given as

$$\mathbf{L}_{p,n}^{(k)} \times \mathbf{M}_{p,n-1}^{(k+1)}\big|_\alpha, \quad \alpha = x, y. \tag{22}$$

This implies that for the kth iteration, information of the $(k+1)$th magnetization state is needed. This is possible in a sequential update where propagations are one step at a time immediately followed by a single control update belonging to the same step. The Krotov approach was introduced to optical spectroscopy by Tannor *et al.*[33] and later Zhu and Rabitz[34] in two different variants. These variants were combined by Maday and Turinici[50] through the introduction of an auxiliary set of controls and two scalar constants, $\delta, \eta \in [0, 2]$ mixing the ordinary and extra controls in the control update. In this representation, Tannor[33] and Zhu-Rabitz[34] formulations became special cases with $\delta = 1, \eta = 0$ and $\delta = 1, \eta = 1$, respectively. Following its introduction to NMR by Maximov *et al.*,[24] we have adopted this generalized formulation in our work and the update procedure reads

$$\omega_{\alpha,n}^{(k+1)} = (1-\delta)\tilde{\omega}_{\alpha,n}^{(k)} + \delta\frac{1}{2\lambda PM_0^2}\sum_{p=1}^{P}\beta_p\mathbf{L}_{p,n}^{(k)}\times\mathbf{M}_{p,n-1}^{(k+1)}\big|_\alpha, \tag{23}$$

$$\tilde{\omega}_{\alpha,n}^{(k+1)} = (1-\eta)\omega_{\alpha,n}^{(k+1)} + \eta\frac{1}{2\lambda PM_0^2}\sum_{p=1}^{P}\beta_p\mathbf{L}_{p,n}^{(k+1)}\times\mathbf{M}_{p,n-1}^{(k+1)}\big|_\alpha. \tag{24}$$

The first iteration with the initial controls is similar to GRAPE, i.e., a full forward and backward propagation. The kth iteration of Krotov involves these steps,

1. Update $\omega_{\alpha,1}^{(k)}$ to $\omega_{\alpha,1}^{(k+1)}$ and propagate $\mathbf{M}_0^{(k+1)}$ to $\mathbf{M}_1^{(k+1)}$
   $\vdots$
   Update $\omega_{\alpha,n}^{(k)}$ to $\omega_{\alpha,n}^{(k+1)}$ and propagate $\mathbf{M}_{n-1}^{(k+1)}$ to $\mathbf{M}_n^{(k+1)}$
   $\vdots$
2. Update $\omega_{\alpha,N}^{(k)}$ to $\omega_{\alpha,N}^{(k+1)}$ and propagate $\mathbf{M}_{N-1}^{(k+1)}$ to $\mathbf{M}_N^{(k+1)}$
3. Update $\tilde{\omega}_{\alpha,N}^{(k)}$ to $\tilde{\omega}_{\alpha,N}^{(k+1)}$ and propagate $\mathbf{L}_N^{(k+1)}$ to $\mathbf{L}_{N-1}^{(k+1)}$
   $\vdots$
   Update $\tilde{\omega}_{\alpha,n}^{(k)}$ to $\tilde{\omega}_{\alpha,n}^{(k+1)}$ and propagate $\mathbf{L}_n^{(k+1)}$ to $\mathbf{L}_{n-1}^{(k+1)}$
   $\vdots$

4. Update $\tilde{\omega}_{\alpha,1}^{(k)}$ to $\tilde{\omega}_{\alpha,1}^{(k+1)}$ and propagate $\mathbf{L}_1^{(k+1)}$ to $\mathbf{L}_0^{(k+1)}$
5. Increment k and go to step 1, unless convergence criterion for cost functional is met.

In the supplementary material,[42] we provide a general proof that the Krotov approach is monotonic, i.e., $J^{(k+1)} - J^{(k)} \geq 0$. The proof demonstrates a theoretical monotonicity for both δ and η in the range of 0 to 2. As will be seen in the presented comparative study, however, this monotonicity can get compromised for certain ill parameters. This fact has been studied in more detail by Maday *et al.*[51] and Maximov *et al.*[29] The iteration loop should in this case stop and the previous set of controls may be extracted. The δ and η parameters replace the role ε has in GRAPE. They can be calibrated independently between 0 and 2. Once a good set of (δ, η) is found, the general Maday-Turinici[50] formulation of the Krotov-based approach is quite robust with respect to different optimizations.

E. GRAPE-BFGS

BFGS is a general, widely used quasi-Newton update method. Newton methods in the search for a stationary point exploit the first- and second-order derivative information, the gradients, and Hessian matrix, respectively. However, it is costly to evaluate the Hessian matrix exactly, and quasi-Newton methods, among which BFGS are very popular, are often used to build up the Hessian iteratively from the gradients. Fouquieres and co-workers[40] found that BFGS in conjunction with GRAPE provides a very efficient OC algorithm for NMR and quantum control problems that is now a standard in the Spinach software.[52,53] The GRAPE-BFGS in this study (and in Spinach) is implemented with MATLAB's fminunc function with explicitly calculated gradients, according to Eq. (14). In our implementation, GRAPE-BFGS receives as an initial guess the controls made with one GRAPE iteration. When the fminunc function calls for a gradient evaluation, the forward and backward propagation is carried out as described for the GRAPE approach. A scalar similar to ε in GRAPE is used to resize the gradient to a level that fminunc finds tolerable. As an element in our optimization software development, we pursued first- as well as second-order finite difference gradients for the GRAPE-BFGS method as proposed by Fouquieres *et al.*[40] By using exact gradients or finite difference approximations instead of the first order approximations as in GRAPE, the scalar ε should be redundant. However, we did not observe any significant improvement with a second-order finite difference gradient, and the gradient calculation time increased 165%–200% depending on the discretization level compared with our implemented first-order gradient. With our focus being *in vivo* MRI applications, we have prioritized keeping the computation time to a minimum. A more extensive explanation of BFGS is given in the supplementary material,[42] Sec. III.

F. Krotov-BFGS

Our implementation of Krotov-BFGS follows the work of Eitan *et al.*,[41] however, we have cast it in the more general Maday-Turinici[50] formulation to provide flexibility to establish the fastest Krotov-based protocol through appropriate choices of the variables δ and η. One difference between the Krotov-BFGS approach relative to the GRAPE or original Krotov approaches is the penalty term in Eq. (3). Eitan *et al.*[41] use a variant that includes the Hessian matrices. In our comparative study, we have instead maintained the penalty term as in Eq. (5) to be able to do a fair comparison. The role of the Hessians in the penalty term as suggested by Eitan *et al.*[41] is still being investigated.

The $N \times N$ Hessian matrices, $\mathbf{H}_\alpha$, $\tilde{\mathbf{H}}_\alpha$ ($\alpha = x, y$) without cross terms, are used in the control updates,

$$
\begin{aligned}
\omega_{\alpha,n}^{(k+1)} = &(1-\delta)\tilde{\omega}_{\alpha,n}^{(k)} \\
&+ \frac{\frac{\delta}{2\lambda M_0^2 P}\sum_{p=1}^{P}\beta_p \mathbf{L}_{p,n}^{(k)} \times \mathbf{M}_{p,n-1}^{(k+1)}|_\alpha}{\mathbf{H}_\alpha^{(k)}(n,n)} \\
&- \delta\frac{\sum_{j=1}^{n-1}\mathbf{H}_\alpha^{(k)}(n,j)\omega_{\alpha,j}^{(k+1)}}{\mathbf{H}_\alpha^{(k)}(n,n)},
\end{aligned} \tag{25}
$$

$$
\begin{aligned}
\tilde{\omega}_{\alpha,n}^{(k+1)} = &(1-\eta)\omega_{\alpha,n}^{(k+1)} \\
&+ \frac{\frac{\eta}{2\lambda M_0^2 P}\sum_{p=1}^{P}\beta_p \mathbf{L}_{p,n}^{(k+1)} \times \mathbf{M}_{p,n-1}^{(k+1)}|_\alpha}{\tilde{\mathbf{H}}_\alpha^{(k)}(n,n)} \\
&- \eta\frac{\sum_{j=1}^{n-1}\tilde{\mathbf{H}}_\alpha^{(k)}(n,j)\tilde{\omega}_{\alpha,j}^{(k+1)}}{\tilde{\mathbf{H}}_\alpha^{(k)}(n,n)},
\end{aligned} \tag{26}
$$

with $(\cdot,\cdot)$ denoting matrix indexing and for $n = 1$, the last term in Eqs. (25) and (26) is not evaluated.

We note that the Hessian matrices, which initially are the $N \times N$ identity matrices, in accord with the original BFGS method also require an update,

$$
\mathbf{H}_\alpha^{(k+1)} = \mathbf{H}_\alpha^{(k)} + \frac{\mathbf{y}_\alpha^{(k)}\mathbf{y}_\alpha^{(k)\mathsf{T}}}{\mathbf{y}_\alpha^{(k)\mathsf{T}}\mathbf{d}_\alpha^{(k)}} - \frac{\mathbf{H}_\alpha^{(k)}\mathbf{d}_\alpha^{(k)}(\mathbf{H}_\alpha^{(k)}\mathbf{d}_\alpha^{(k)})^\mathsf{T}}{\mathbf{d}_\alpha^\mathsf{T}\mathbf{H}_\alpha^{(k)}\mathbf{d}_\alpha^{(k)}}, \tag{27}
$$

$$
\tilde{\mathbf{H}}_\alpha^{(k+1)} = \tilde{\mathbf{H}}_\alpha^{(k)} + \frac{\tilde{\mathbf{y}}_\alpha^{(k)}\tilde{\mathbf{y}}_\alpha^{(k)\mathsf{T}}}{\tilde{\mathbf{y}}_\alpha^{(k)\mathsf{T}}\tilde{\mathbf{d}}_\alpha^{(k)}} - \frac{\tilde{\mathbf{H}}_\alpha^{(k)}\tilde{\mathbf{d}}_\alpha^{(k)}(\tilde{\mathbf{H}}_\alpha^{(k)}\tilde{\mathbf{d}}_\alpha^{(k)})^\mathsf{T}}{\tilde{\mathbf{d}}_\alpha^\mathsf{T}\tilde{\mathbf{H}}_\alpha^{(k)}\tilde{\mathbf{d}}_\alpha^{(k)}}, \tag{28}
$$

with

$$
\mathbf{d}_\alpha^{(k)} = \vec{\omega}_\alpha^{(k+1)} - \vec{\omega}_\alpha^{(k)}, \tag{29}
$$

$$
\tilde{\mathbf{d}}_\alpha^{(k)} = \vec{\tilde{\omega}}_\alpha^{(k+1)} - \vec{\tilde{\omega}}_\alpha^{(k)}, \tag{30}
$$

$$
\mathbf{y}_\alpha^{(k)} = \nabla J_\alpha^{(k+1)} - \nabla J_\alpha^{(k)}, \tag{31}
$$

$$
\tilde{\mathbf{y}}_\alpha^{(k)} = \tilde{\nabla} J_\alpha^{(k+1)} - \tilde{\nabla} J_\alpha^{(k)}. \tag{32}
$$

Here $\vec{\omega}_\alpha$ and $\vec{\tilde{\omega}}_\alpha$ in Eqs. (29) and (30), respectively, constitute the entire arrays of controls from 1 to N for the kth iteration. These arrays are in this respect considered as column vectors. As proposed by Eitan *el al.*,[41] the line search is skipped, and therefore, $\mathbf{d}_q$ and $\tilde{\mathbf{d}}_q$ are not scaled by any factor. The ∇J_α and

$\tilde{\nabla} J_\alpha$ in Eqs. (31) and (32), respectively, are the control gradients as in the usual GRAPE method with the nth component given by Eq. (14).

The Krotov-BFGS follows the iterative scheme of Krotov except for the update procedure now given by Eqs. (25) and (26). Between the steps 4 and 5, the Hessian matrices are updated using Eqs. (27) and (28). In other words, the Hessians are not updated along the way as the controls are, which, on the other hand, can only exploit the Hessian information up to the given index of control update. Specifically, the derivations of Eqs. (25) and (26), and further details of Eqs. (27) and (28) are given in the supplementary material,[42] Sec. III.

III. COMPARATIVE ANALYSIS

With attention to MRI applications, the three algorithms Krotov, Krotov-BFGS, and GRAPE-BFGS were presented to a series of performance tests. We shall leave the theoretical well-established MRI details for MDRF pulses to the original papers on this matter[14,54] and only provide a summary of the technical details. The spatial profile which the MDRF pulse is targeted for can be expressed in a spatial frequency domain by a Fourier transform. This spatial frequency space is often called the excitation k-space and the field gradient waveform that accompanies the MDRF pulse determines to which locations in the excitation k-space that rf energy is deposited. To satisfy the general sampling criteria, the field gradients must perform a sufficiently dense excitation k-space traversal. A spiral trajectory ending in the excitation k-space center is often the choice since it is area effective and provides a coherent magnetization at the end of the pulse. The excitation k-space traversal was accomplished using a spiral of 16 revolutions yielding a resolution of 1.8 mm for a field-of-excitation (FOX) of 56 mm. The gradient waveform accommodating this spiral trajectory was established using the slew-rate- and amplitude-limited design by Glover.[55] In this design, the gyromagnetic ratio γ, the maximum slew-rate (2000 T/m/s), the maximum amplitude (0.4 T/m), the FOX, and the resolution dictate the pulse duration, which in the present case amounts to 3.9 ms. The MDRF pulses being optimized were targeted for creating a 90° flip-angle of $\mathbf{M} = [0, 0, M_0]^\mathsf{T}$ to $\mathbf{M} = [M_0, 0, 0]^\mathsf{T}$ inside the region-of-interest (ROI) and leave the area outside unperturbed. In order to avoid any bias associated with the shape of the ROI, the algorithms were tested on 100 different ROI's that were generated on basis of four shapes (bar, triangle, circle, and half circle) mixed together with random rotation and translocation, see Figs. 1 and 2 in Sec. IV A in the supplementary material.[42] All ROI's were described in a 64×64 spatial grid.

A. Parameter optimization

The algorithms have a few parameters that calibrate the convergence properties; δ and η for Krotov and Krotov-BFGS, and ε for GRAPE-BFGS. Common to all three methods are the parameters λ and Δt. A search for good parameters was accomplished prior to the actual comparison. Of the 100

randomly generated ROI's, only 10 were used for the parameter space search. All three algorithms were tested with $\lambda [\mu s]$ = 0.0001, 0.001, 0.01, 0.05, 0.1, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, and 9. Krotov and Krotov-BFGS were tested with δ and η ranging from 0 to 2 in steps of 0.125. GRAPE-BFGS was tested with $\varepsilon[s^{-1}] = 1 \times 10^{-2}$, 5×10^{-4}, 1×10^{-4}, 5×10^{-5}, 1×10^{-5}, 5×10^{-6}, and 1×10^{-6}.

Three random initial guesses (see Sec. IV B in the supplementary material[42] for details) were constructed with Δt = 2, 4, and 8 μs to test the performance in terms of pulse discretization. The same initial guess for a particular Δt was used for all three algorithms, while all other parameters were subject to variation. Every optimization was allowed to run 9 iterations, which in our experience is sufficient for the purpose identifying "good" parameters. The full set of data resulting from the parameter search are presented in Sec. IV C of the supplementary material.[42]

Figure 1 provides an example of a parameter optimization for a $\Delta t = 2\,\mu s$ pulse sequence in terms of the total cost J, the final cost Φ, and the running cost Ξ. For Krotov and Krotov-BFGS, since δ and η can be varied independently, it is relevant to display the data as $J(\delta, \eta)$-, $\Phi(\delta, \eta)$-, and $\Xi(\delta, \eta)/\lambda$-maps. This is exemplified in Fig. 1 for the case of $\lambda = 0.5\,\mu s$. For GRAPE-BFGS, we displayed the corresponding (λ, ε)-maps. Two contours on top of the $J(\delta, \eta)$-maps divide it into sections having complete monotonic convergence throughout
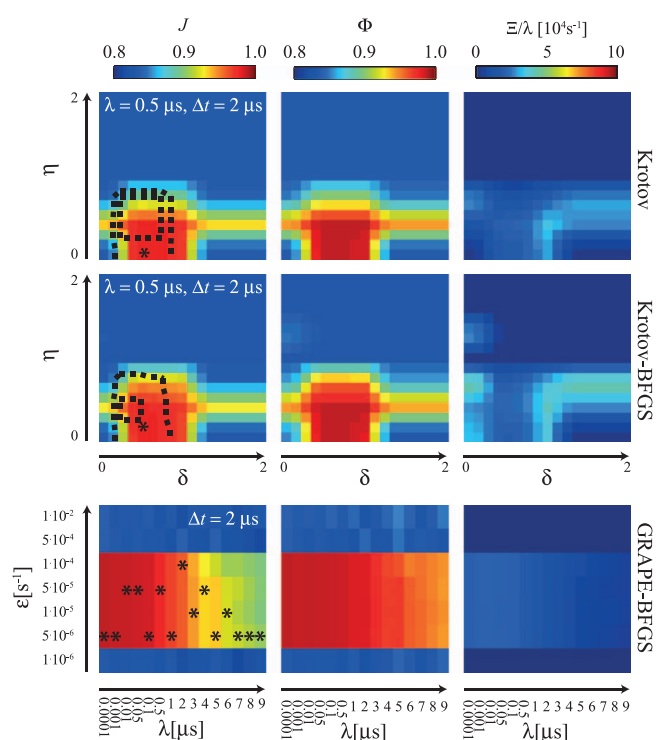

FIG. 1. Example of the parameter space optimization. The top and central panels display the $J(\delta, \eta)$-, $\Phi(\delta, \eta)$-, and $\Xi(\delta, \eta)/\lambda$-maps with $\Delta t = 2\,\mu s$ and $\lambda = 0.5\,\mu s$ for Krotov and Krotov-BFGS, respectively. The black *'s signify the maximum value and the dashed contours divide the maps into sections having complete monotonic convergence throughout the 9 iterations (inside inner contour), lost monotonicity after 5 iterations (between contours), and monotonicity in fewer than 5 iterations (outside outer contour). The bottom panel displays the $J(\lambda, \varepsilon)$-, $\Phi(\lambda, \varepsilon)$-, and $\Xi(\lambda, \varepsilon)/\lambda$-maps with $\Delta t = 2\,\mu s$ for GRAPE-BFGS. The black *'s show the maxima for a given λ.

the 9 iterations (innermost), lost monotonicity after 5 iterations (middle), and monotonicity of fewer than 5 iterations. The maxima are displayed with *'s and are seen to lie outside the regions of complete monotonicity. For many repeated optimizations as in this comparative study, it may therefore be necessary to choose "stable" monotonic parameters rather than "maximizing" parameters.

Typically δ and η should preferably be lower than 0.5 for both Krotov and Krotov-BFGS. The monotonicity areas for Krotov-BFGS are smaller than for Krotov. Maday and Turinici[50] found that $(\delta, \eta) = (2, 0), (5/3, 0), (4/3, 0), (1, 0), (5/3, 1/3)$, and $(2, 1/3)$ are good parameters in one of their studies. Maximov *et al.*[24] had good convergence in the same corner of parameters as Krotov and Krotov-BFGS, yet their algorithm covers a larger area. It is interesting to note that neither $(1, 0)$ as in the Tannor approach,[33] nor $(1, 1)$ as in the Zhu-Rabitz approach[34] seems to work in the present application. Supplementary tests (not shown) have been concentrated around the (δ, η) regions $(1, 0)$ and $(1, 1)$, however, with a wider and higher digitized λ-range. Neither of these searches provided useful λ-values for MDRF pulses.

With $\Delta t = 2\,\mu s$, GRAPE-BFGS displays quite a broad range of suitable ε-values; $\varepsilon \in [1 \times 10^{-4}, \ldots, 5 \times 10^{-6}]\,s^{-1}$. The *'s highlighting the maxima are virtually redundant as there is no significant difference in the ε-range for a given λ. GRAPE-BFGS can in fact optimize without the penalty term, which is also seen for λ values in the negligible range, say, 10^{-10} s. However, in such cases the integrated rf pulse power increases markedly. The color coding used in Fig. 1 is directly comparable to the color coding in Figs. 5 and 6 in the supplementary material that show the entire data set.[42]

Based on this analysis, $\lambda = 2\,\mu s$ was chosen for the common penalty scalar as it provides a good compromise between stable convergence and adequate efficiency. For Krotov and Krotov-BFGS $\delta = \eta = 0.375$ and for GRAPE-BFGS $\varepsilon = 5 \times 10^{-5}\,s^{-1}$, were considered good parameters for a broad range of Δt and the given λ.

B. Convergence analysis

With calibrated optimization parameters, the three optimal control methods were tested for the 100 randomized ROI's using pulse discretizations of $\Delta t = 1, 2, 3, 4, 6$, and 8 μs. To avoid bias towards the initial guess, the results were averaged over 20 different initial guesses with the number of iterations limited to 19. Each algorithm experienced the same 20 initial guesses. To obtain a reliable measure for the time-per-iteration (TPI) for each algorithm, 20 of the 100 cases were computed on a standard Macintosh 2.2 GHz Intel i7 processor with 8 GB RAM, with careful attention to occupy the computer with a minimum of background processes and purged memory etc. For the remaining cases multiple processes were allowed to take advantage of parallel computation.

The TPI's for the three optimal control methods are compared in Fig. 2. It is evident that the TPI's differ substantially in all cases with Krotov being the fastest, e.g., with TPI($\Delta t = 8\,\mu s$) = 0.76 $\pm$ 0.03 s compared to TPI($\Delta t = 8\,\mu s$) = 0.87 $\pm$ 0.06 s and TPI($\Delta t = 8\,\mu s$) = 2.1 $\pm$ 0.39 s for Krotov-BFGS
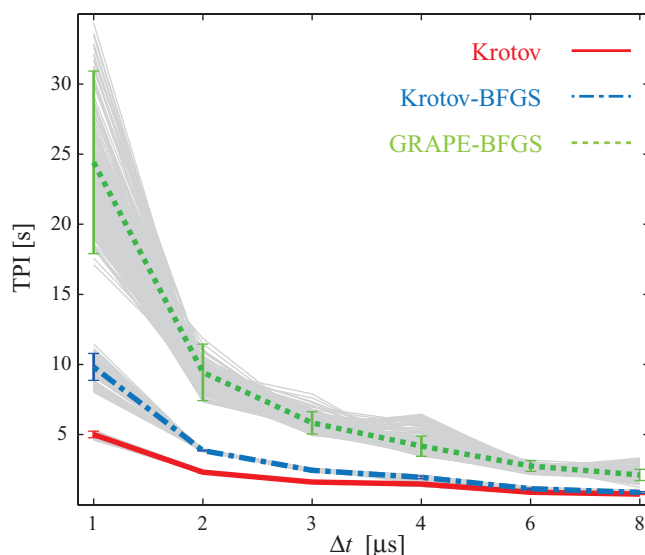

FIG. 2. Time-per-iteration (TPI) for Krotov (red, full), Krotov-BFGS (blue, dashed-dotted), and GRAPE-BFGS (green, dotted) optimal control calculations displayed as a function of the subpulse length Δt. The underlying gray curves illustrate the individual data that the colored curves were averaged upon.

and GRAPE-BFGS, respectively. With respect to the required number of iterations, we found the spread in convergence behavior to be more dependent on the ROI than on the 20 different starting pulse sequences. Accordingly, we applied the convergence analysis to each individual ROI. The so-called *Wilcoxon Signed Rank Test*[56] was used to identify cases between the algorithms that were significantly different based on the 20 paired trials. We considered the data iteration by iteration and Fig. 3 shows which algorithm dominates most cases for a given iteration. Considering the left column of Krotov vs. GRAPE-BFGS, we see that Krotov completely dominates in the first two iterations for any Δt. GRAPE-BFGS is best in the third iteration except when $\Delta t = 1\,\mu s$. The white space above the bars indicates a tie between the algorithms, and in the later iterations white space also accounts for the possibility of stopped optimization of one or both algorithms. It should be noted that Krotov in most of the cases used all 19 iterations and only rarely lost monotonicity. White space in the later iterations is therefore a sign of ended optimization of the other two. GRAPE-BFGS was experienced to escape the optimization after roughly 10 iterations proclaiming inability to proceed. The longer the Δt, the more iterations will GRAPE-BFGS need before converging due to the inaccuracy of the gradients presumably. The same problem applies to Krotov-BFGS that is seen to be most stable for Δt up to 3 μs. The right column reveals, however, that under stable conditions it is not better than Krotov. Only in iteration three is Krotov-BFGS for some undisclosed reason dominating for $\Delta t = 1\,\mu s$.

The TPI data were used to analyse the convergence behavior of algorithms computation-time-wise as opposed to the iteration-wise as in Fig. 3. Curves for the mean values of J versus time are shown in Fig. 4. The vertical errorbars show the standard deviation, which we could conclude were mainly ROI dependent. The horizontal errorbars of standard deviation correspond to the errorbars in Fig. 2. The
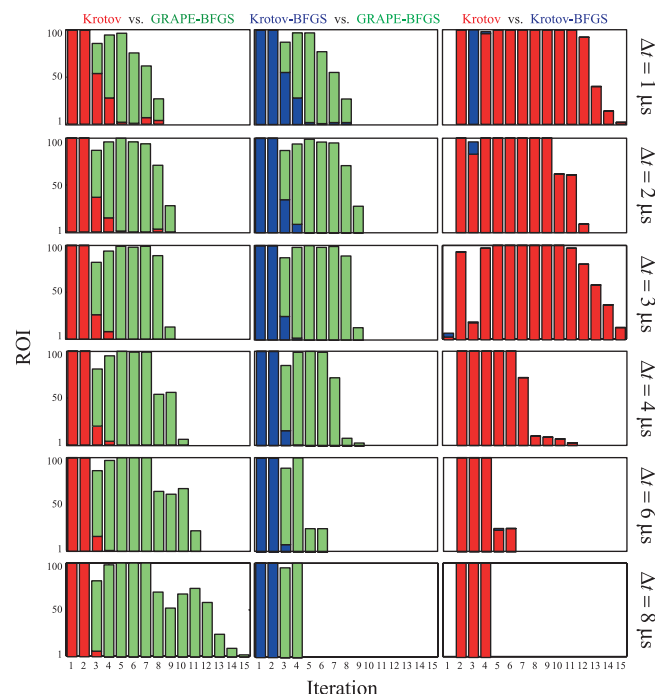
FIG. 3. Barplot illustrating the pair-wise performance of the algorithms Krotov (red), Krotov-BFGS (blue), and GRAPE-BFGS (green) in obtaining highest convergence (J) versus the iteration number for the 100 ROI's investigated. The total bar-height is 100 corresponding to the 100 ROI's, and the number of iterations goes up to 15. White space corresponds to a tie between the algorithms or that at least one algorithm did not provide data for the current iteration. The data for this plot was obtained from the analysis in Fig. 7 in the supplementary material.[42]
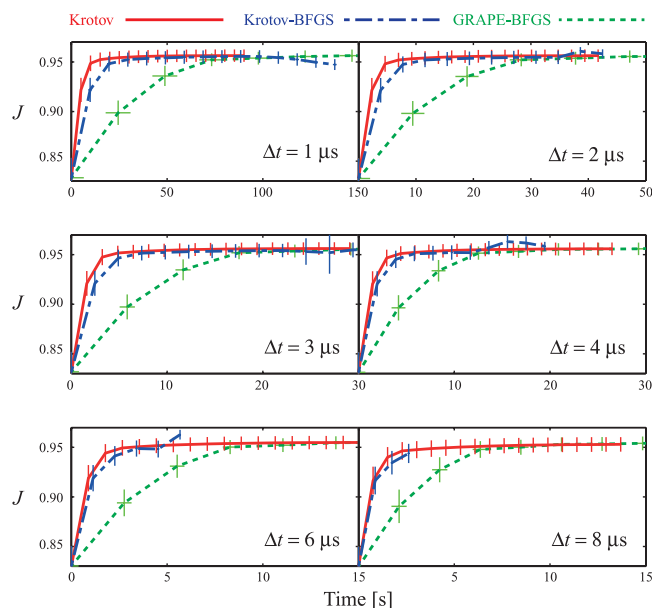

FIG. 4. Convergence of the three algorithms, Krotov (red), Krotov-BFGS (blue, dashed-dotted), and GRAPE-BFGS (green, dotted) vs. time. The curves are the average over the 20 trials and 100 ROIs.

convergence to a level just above 0.95 is due to the running cost term. By lowering λ, this level can be increased with the risk of more instability. The significant TPI differences are again seen here parting the Krotov-based methods from GRAPE-BFGS. As it is the mean curve that is depicted in Fig. 4, the fluctuation of Krotov-BFGS is not because of non-monotonic behavior. Those optimizations that did loose monotonicity were not counted in the mean value. Rather, the fluctuation is a sign of growing numerical instability that leads to loss of monotonicity.

Interestingly, we found that all algorithms agreed almost completely in pin-pointing which of the ROI's were hard or easy to optimize a MDRF sequence for. That is, by sorting all the optimized cases after the highest value J of the cost functional, the indexing was nearly the same for all algorithms. No correlation was found between that order and multiple sorted combinations of ROI specific features, including total area, number of elements, number of corners, and lengths of boundaries.

IV. EXPERIMENT

For the purpose of demonstration, we used our Krotov-based method to design a MDRF that excites the letters "JCP" (letter height 4.6 mm) in a $CuSO_4$-doped (20 mm diameter) water phantom placed in our 16.4 T vertical bore Bruker Avance II spectrometer. The gradient waveform performed a spiral of 25 revolutions for a FOX of 20 mm yielding a reso-

lution of 0.4 mm. The slew-rate was restricted to 3272 T/m/s and amplitude to 722 T/m, which is 30% and 50% of the hardware limits, respectively. The pulse duration was 8.87 ms and the time-discretization was $4\,\mu$s, and we used a random initial guess limited to 10% of the maximum rf field strength (10 kHz). Figure 5(a) shows the optimized MDRF pulse amplitude (ν_1) and phase (ϕ_1). The optimization was performed on basis of a B_0 map, or $\Delta\omega_0(\mathbf{r})$, to compensate off-resonance effects. The B_0 map was obtained in the following way: Two standard gradient echo images were acquired with echo-times (T_E) of 4.0 ms and 4.5 ms, respectively. The gradient echo images were recorded with a FOV of 20 mm with 128 phase-encoding (PE) steps and 128 read-out (RO) steps, i.e., a grid of 128×128, a repetition time (T_R) of 1 s, a slice thickness (TH) of 1 mm, and with the same slice position as the MDRF pulse was intended for. In this case, the slice was axial and
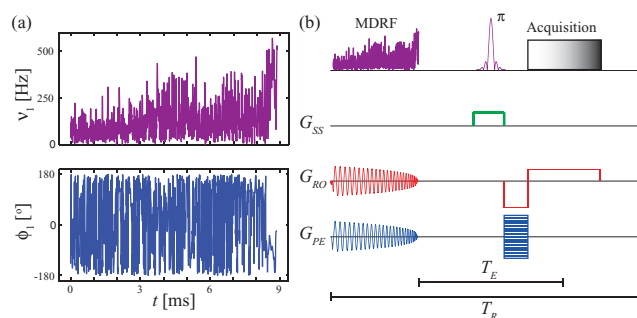

FIG. 5. (a) Schematic representation of the optimized MDRF pulse reported in terms of amplitude (ν_1) and phase (ϕ_1) as being implemented into a standard spin-echo sequence (b). The gradients G_{SS}, G_{RO}, and G_{PE} corresponding to slice-selection, read-out, and phase-encoding, respectively, are not to scale. A signal echo is formed in the center of the acquisition windows at time T_E counted from the end of the MDRF pulse. The PE ladder block is stepped at each repetition. The total experiment time is the number of averages times the number of PE steps times the T_R; 17 minutes in the case shown in Fig. 6.
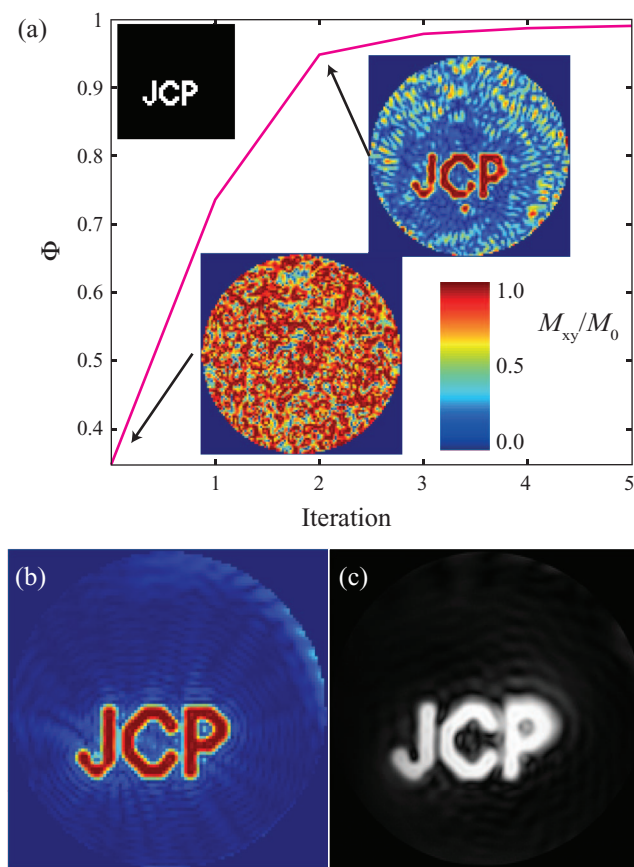
FIG. 6. (a) Development of the final cost, Φ, during the first five iterations of the Krotov-based optimal control design of the MDRF pulse. The black and white insert shows the desired ROI. The two simulated excitation profiles of $|M_{xy}|/M_0$, where $M_{xy} = M_x + iM_y$ in the rotating reference frame, correspond to the initial guess and the second iteration demonstrating how fast the Krotov-based method progresses towards a good solution. (b) Simulation of the final MDRF pulse excitation profile as used in the experimental demonstration. (c) Experimental demonstration of the performance for a optimal control MDRF pulse obtained for CuSO$_4$-doped water phantom.

positioned in the magnet isocenter. The phase-data were subtracted and processed to give a B_0 map as described in Ref. 57, and the B_0 map was included in the optimization by the term $\Delta\omega_0(\mathbf{r})$ in Eq. (2).

The MDRF pulse, resulting from the OC optimization (51 s optimization time), was implemented into a standard spin-echo imaging sequence as seen in Fig. 5(b). The π-pulse refocusses the signal and provides the slice-selection (SS) of the third spatial dimension. The image was acquired with FOV of 20 mm on a 256×256 grid, with a T_E of 14 ms, a T_R of 1 s, and was averaged over four scans. The TH was 1 mm. Figure 6(a) shows a graph of the efficiency progress for the first five iterations in the OC optimization. The ROI is shown in the upper left corner of Fig. 6(a). Two simulations show transverse magnetization profiles of the initial guess and the pulse of iteration two. Figure 6(b) shows the simulation of the 19th iteration that was used for the experimental demonstration in Fig. 6(c) in the CuSO$_4$-doped water phantom.

To demonstrate the general applicability of optimal control for MDRF pulses in MRI using commonly available simulation software, our second example addresses the use of the GRAPE algorithm as implemented in the open-source SIMP-
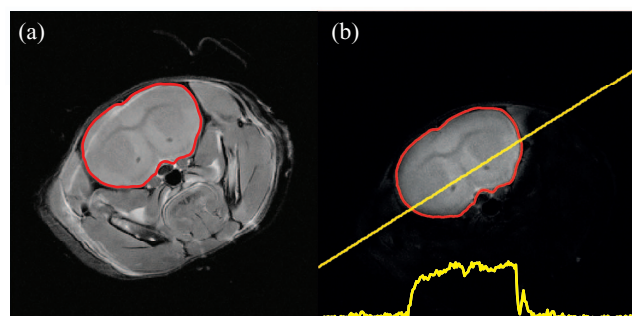

FIG. 7. (a) MRI prescan of the decapitated mouse head. The red loop designates the desired ROI. (b) The result from a MDRF pulse designed using GRAPE as implemented into the SIMPSON software package.[23, 49] The yellow profile is taken along the yellow straight line.

SON software.[23, 49] The MDRF pulse was targeted for exciting the brain in an *ex vivo* mouse. Figure 7(a) shows a prescan of the mouse head with the red loop indicating the ROI boundary. The gradient waveform performed a spiral of 17.5 revolutions for a FOX of 20 mm yielding a resolution of 0.571 mm. The slew-rate and amplitude were in this case limited to 50% and 100% of the hardware limits, respectively. The pulse duration was 4.53 ms and the time-discretization was 8 μs. The actual gradient waveform was measured using the approach described by Alley *et al.*[58] and replaced the nominal gradient waveform in the optimization. With the same parameters as above a B_0 map was acquired and included in the optimization. On basis of the ROI, the actual gradient waveform and this B_0 map, we obtained an MDRF pulse, which led to the result in Fig. 7(b). The yellow profile in the bottom corresponds to the trace of the yellow straight line.

V. CONCLUSION

In the conclusion, we have demonstrated the use of optimal control in MRI with attention to fast design of advanced multidimensional rf pulses. The Krotov-based optimal control approach was adapted to the classical Bloch description of magnetization and various means to advance the MATLAB scripts for faster computation were tested. The various scripts will be released on our website upon publication.[59] To establish the fastest protocol, being immensely important for *in vivo* applications, we performed a detailed comparative analysis of the proposed Krotov algorithm with a second-order gradient-based method (GRAPE-BFGS) and a hybrid scheme of the two, Krotov-BFGS. On basis of this analysis, we found our Krotov-based approach to be nearly three and five times faster than GRAPE-BFGS algorithm using a course and fine discretization, respectively. Also, the Krotov approach appear more efficient than the Krotov-BFGS approach. It was shown that the Krotov methods need only 2–3 iterations to find a decent MDRF pulse sequence without the need for a good initial guess, as opposed to the gradient methods being more heavily dependent on such input. The optimal control strategy indeed proves well suited for fast design of high-quality MDRF pulses, which we demonstrated by the excitation of an advanced symbol pattern in a water phantom using the Krotov-based approach as well as a mouse brain using the GRAPE

approach available in SIMPSON. We anticipate that our introduction of fast Krotov-based optimal control in a Bloch equation formulation will have a significant impact on future applications of numerical design of advanced experiments within *in vivo* MRI and MRS. This applies to medical applications calling for fast design procedures, but certainly also the MR based microscopy in wider areas of science, including, for example, materials science and nanomedicine.

ACKNOWLEDGMENTS

[1]M. Wallentin, A. H. Nielsen, P. Vuust, A. Dohn, A. Roepstorff, and T. E. Lund, NeuroImage **58**, 963 (2011).

[2]S. A. Butt, L. V. Søgaard, P. O. Magnusson, M. H. Lauritzen, C. Laustsen, P. Åkeson, and J. H. Ardenkjær-Larsen, J. Cereb. Blood. Flow. Metab. (2012).

[3]J. J. Flint, C. H. Lee, B. Hansen, M. Fey, D. Schmidig, J. D. Bui, M. A. King, P. Vestergaard-Poulsen, and S. J. Blackband, NeuroImage **46**, 1037 (2009).

[4]S. N. Jespersen, C. R. Bjarkam, J. R. Nyengaard, M. M. Chakravarty, B. Hansen, T. Vosegaard, L. Østergaard, D. Yablonskiy, N. C. Nielsen, and P. Vestergaard-Poulsen, NeuroImage **49**, 205 (2010).

[5]P. Vestergaard-Poulsen, G. Wegener, B. Hansen, C. R. Bjarkam, S. J. Blackband, N. C. Nielsen, and S. N. Jespersen, PLoS ONE **6**, e20653 (2011).

[6]B. Hansen, J. J. Flint, C. Heon-Lee, M. Fey, F. Vincent, M. A. King, P. Vestergaard-Poulsen, and S. J. Blackband, NeuroImage **57**, 1458 (2011).

[7]J. Flint, B. Hansen, S. Portnoy, C.-H. Lee, M. A. King, M. Fey, F. Vincent, G. J. Stanisz, P. Vestergaard-Poulsen, and S. J. Blackband, NeuroImage **60**, 1404 (2012).

[8]P.-M. L. Robitaille, A. M. Abduljalil, and A. Kangarlu, J. Comput. Assist. Tomo. **24**, 2 (2000).

[9]T. Vaughan, L. DelaBarre, C. Snyder, J. Tian, C. Akgun, D. Shrivastava, W. Liu, C. Olson, G. Adriany, J. Strupp, P. Andersen, A. Gopinath, P.-F. van de Moortele, M. Garwood, and K. Ugurbil, Magn. Reson. Med. **56**, 1274 (2006).

[10]H. P. Hetherington, J. W. Pan, W.-J. Chu, G. F. Mason, and B. R. Newcomer, NMR Biomed. **10**, 360 (1997).

[11]A. C. Silva and A. P. Koretsky, Proc. Natl. Acad. Sci. U.S.A. **99**, 15182 (2002).

[12]U. Katscher, P. Börnert, C. Leussler, and J. S. van den Brink, Magn. Reson. Med. **49**, 144 (2003).

[13]Y. Zhu, Magn. Reson. Med. **51**, 775 (2004).

[14]J. M. Pauly, D. G. Nishimura, and A. Macovski, J. Magn. Reson. **81**, 43 (1989).

[15]C.-y. Yip, J. A. Fessler, and D. C. Noll, Magn. Reson. Med. **54**, 908 (2005).

[16]C.-y. Yip, J. A. Fessler, and D. C. Noll, Magn. Reson. Med. **56**, 1050 (2006).

[17]L. Pontryagin, B. Boltyanskii, R. Gamkrelidze, and E. Mishchenko, *The Mathematical Theory of Optimal Processes* (Wiley-Interscience, New York, 1962).

[18]A. E. Bryson and Y.-C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control* (Hemisphere Publishing Corporation, Washington, DC, 1975).

[19]V. F. Krotov, *Global Methods in Optimal Control Theory* (Marcel Dekker, New York, 1995).

[20]N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbruggen, and S. J. Glaser, J. Magn. Reson. **172**, 296 (2005).

[21]T. E. Skinner, N. I. Gershenzon, M. Nimbalkar, W. Bermel, B. Luy, and S. J. Glaser, J. Magn. Reson. **216**, 78 (2012).

[22]C. T. Kehlet, A. C. Sivertsen, M. Bjerring, T. O. Reiss, N. Khaneja, S. J. Glaser, and N. C. Nielsen, J. Am. Chem. Soc. **126**, 10202 (2004).

[23]Z. Tošner, T. Vosegaard, C. Kehlet, N. Khaneja, S. J. Glaser, and N. C. Nielsen, J. Magn. Reson. **197**, 120 (2009).

[24]I. I. Maximov, Z. Tošner, and N. C. Nielsen, J. Chem. Phys. **128**, 184505 (2008).

[25]C. Bretschneider, A. Karabanov, N. C. Nielsen, and W. Köckenberger, Chem. Phys. **136**, 094301 (2012).

[26]S. M. Conolly, D. G. Nishimura, and A. Macovski, IEEE Trans. Med. Imag. **5**, 106 (1986).

[27]D. Xu, K. F. King, Y. Zhu, G. C. McKinnon, and Z.-P. Liang, Magn. Reson. Med. **59**, 547 (2008).

[28]J. Li, J. Ruths, T. Yu, H. Arthanari, and G. Wagner, Proc. Natl. Acad. Sci. U.S.A. **108**, 1879 (2011).

[29]I. I. Maximov, J. Salomon, G. Turinici, and N. C. Nielsen, J. Chem. Phys. **132**, 084107 (2010).

[30]A. Sbrizzi, H. Hoogduin, J. J. Lagendijk, P. Luijten, G. L. G. Sleijpen, and C. A. T. van den Berg, Magn. Reson. Med. **66**, 879 (2011).

[31]W. A. Grissom, D. Xu, A. B. Kerr, J. A. Fessler, and D. C. Noll, IEEE Trans. Med. Imag. **28**, 1548 (2009).

[32]M. R. Hestenes and E. Stiefel, J. Res. Natl. Bur. Stand. **49**, 409 (1952).

[33]D. J. Tannor, V. Kazakov, and V. Orlov, *Control of Photochemical Branching: Novel Procedures for Finding Optimal Pulses and Global Upper Bounds* (Plenum, New York, 1992).

[34]W. Zhu and H. Rabitz, J. Chem. Phys. **109**, 385 (1998).

[35]D. M. Reich, M. Ndong, and C. P. Koch, J. Chem. Phys. **136**, 104103 (2012).

[36]C. Broyden, J. Inst. Math. Appl. **6**, 76 (1970).

[37]R. Fletcher, Comput. J. **13**, 317 (1970).

[38]D. Goldfarb, Math. Comput. **24**, 23 (1970).

[39]D. Shanno, Math. Comput. **24**, 647 (1970).

[40]P. de Fouquieres, S. G. Schirmer, S. J. Glaser, and I. Kuprov, J. Magn. Reson. **212**, 412 (2011).

[41]R. Eitan, M. Mundt, and D. J. Tannor, Phys. Rev. A **83**, 053426 (2011).

[42]See supplementary material at http://dx.doi.org/10.1063/1.4739755 for numerical details, proof of monotonicity regarding the Krotov-based algorithms, and supporting data for the comparative study.

[43]E. K. Insko and L. Bolinger, J. Magn. Reson., Ser. A **103**, 82 (1993).

[44]A. Maudsley, H. Simon, and S. Hilal, J. Phys. E **17**, 216 (1984).

[45]MATLAB Version R2011b, The MathWorks Inc., Natick, MA, 2011.

[46]C. Kehlet, M. Bjerring, A. C. Sivertsen, T. Kristensen, J. J. Enghild, S. J. Glaser, N. Khaneja, and N. C. Nielsen, J. Magn. Reson. **188**, 216 (2007).

[47]G. B. Matson, K. Young, and L. G. Kaiser, J. Magn. Reson. **199**, 30 (2009).

[48]D. Rosenfeld and Y. Zur, Magn. Reson. Med. **36**, 401 (1996).

[49]M. Bak, J. T. Rasmussen, and N. C. Nielsen, J. Magn. Reson. **147**, 296 (2000).

[50]Y. Maday and G. Turinici, J. Chem. Phys. **118**, 8191 (2003).

[51]Y. Maday, J. Salomon, and G. Turinici, Numer. Math. **103**, 323 (2006).

[52]H. J. Hogben, M. Krzystyniak, G. Charnock, P. J. Hore, and I. Kuprov, J. Magn. Reson. **208**, 179 (2011).

[53]Spinach Version 1.2.1217, Theoretical Spin Dynamics Group, University of Southampton, 2012.

[54]P. Börnert and B. Aldefeld, MAGMA **7**, 166 (1998).

[55]G. Glover, Magn. Reson. Med. **42**, 412 (1999).

[56]F. Wilcoxon, Biometr. Bull. **1**, 80 (1945).

[57]M. A. Bernstein, K. F. King, and X. J. Zhou, *Handbook of MRI Pulse Sequences* (Elsevier Academic, 2004).

[58]M. Alley, G. Glover, and N. Pelc, Magn. Reson. Med. **39**, 581 (1998).

[59]See www.bionmr.chem.au.dk for MATLAB scripts for optimal control design of MDRF.